

# Boas Práticas na Migração de Repositórios: lições aprendidas com o CALM e o ARQBASE

*José Carlos Ramalho, Miguel Ferreira, Luís Faria, Rui Castro*

KEEP SOLUTIONS

Rua Rosalvo de Almeida

4710-429 Braga

Tel: +351 253066735

{jcr, mferreira, lfaria, rcastro}@keep.pt

## RESUMO

Atualmente, na administração pública portuguesa, o suporte digital é incontornável. Nos últimos anos, os registos em papel têm vindo a ser gradualmente substituídos por registos digitais dando origem a consideráveis sistemas de informação ou repositórios digitais.

Quem trabalha com um sistema de informação, ao fim de algum tempo, dá como certa a sua longevidade, segurança, e capacidade de conservação da informação que armazena. No entanto, a realidade não é bem essa, havendo diversas razões que nos podem levar a uma migração do sistema de informação, como por exemplo, a obsolescência da plataforma tecnológica da qual o sistema de informação depende para funcionar, o sistema já não satisfaz os requisitos da comunidade de utilizadores, o sistema perdeu o suporte ou manutenção (os fornecedores já não fornecem o serviço ou o contrato caducou e não foi renovado), redução de custos (passou a haver soluções tecnologicamente mais evoluídas e mais compensadoras em termos de custos), não suporta interoperabilidade (atualmente, este requisito tem vindo a assumir uma importância cada vez maior).

Seja o repositório de uma biblioteca, de um arquivo, ou de uma empresa, chegará um dia em que os dados geridos por um determinado sistema de informação terão de ser migrados para um novo repositório de dados.

Este artigo pretende ser um guia de boas práticas no suporte à migração de sistemas legados: irão descrever-se as diferentes fases do processo, atividades e riscos associados. Este guia resulta de uma combinação de 13 metodologias existentes que foram analisadas, otimizadas e combinadas.

**PALAVRAS-CHAVE:** Migração de Repositórios, Detecção e relato de anomalias, CALM, ARQBASE

## INTRODUÇÃO

Muitas instituições são proprietárias de sistemas que suportam repositórios digitais de longa duração que se encontram em operação há vários anos. Muitos destes sistemas estão a chegar ao fim da sua vida e serão substituídos pela próxima geração de sistemas digitais.

Esta situação implicará a migração de toda a informação: milhões de ficheiros compreendendo vários

terabytes de informação. Devido à dimensão desta operação e ao impacto que poderá causar, este processo deve ser planeado, validado e executado cuidadosamente.

Podem ser várias e variadas as razões pelas quais uma organização decide migrar o seu repositório digital. Por exemplo:

- O sistema já não responde às necessidades atuais: não tem algumas características, baixo desempenho, baixa capacidade, interoperabilidade, usabilidade, e outros.;
- Cortes orçamentais ditam a adoção de uma solução mais económica;
- O Fornecedor responsável pela manutenção do sistema deixou de existir;
- O Fornecedor não tem um nível de serviço de suporte satisfatório;
- O ambiente tecnológico do sistema deixou de ser suportado (e.g. atualizações de segurança e correção de erros).

Como exemplos de projetos de migração podemos considerar os seguintes cenários [1]:

- Migração de um sistema relativamente simples para outro;
- Atualização da versão de um sistema;
- Migrar vários sistemas convergindo a sua informação num sistema novo;
- Migração de um sistema crítico em que a migração deverá ocorrer sem que o sistema seja interrompido;
- Migração de vários sistemas concorrentes com consolidação da informação.

Pode parecer que dois sistemas que mantêm o mesmo tipo de informação e que realizam o mesmo tipo de tarefas são semelhantes e que por isso a migração entre um e outro seja simples. No entanto, este pressuposto está longe da realidade. Por exemplo, os sistemas legados não possuem invariantes sobre a integridade dos dados [6]. Campos que deviam ser povoados por valores oriundos de uma lista de valores pré-definida não o são, levando a que apareçam valores “estranhos” que têm de ser alvo de especial atenção durante a migração.

Outro problema comum está relacionado com as diferenças entre os modelos hierárquico e relacional. Os sistemas hierárquicos não respeitam as regras de normalização e permitem o armazenamento redundante de informação. Há uns anos atrás utilizaram-se sistemas de gestão de bases de dados baseados no modelo hierárquico (ARQBASE) mas hoje o modelo relacional domina as soluções existentes por isso surge, muitas vezes, a necessidade de migrar de um para o outro. [6]

Mas podemos encontrar outras dificuldades na migração de um para outro:

- A extração de informação do sistema legado pode ser extremamente complexa: a funcionalidade de exportação não existe, o suporte técnico não está disponível, a documentação é incompleta ou inexistente, a organização não possui as credenciais necessárias para ter acesso aos dados;
- O mapeamento estrutural entre os modelos dos repositório pode ser difícil ou mesmo impossível (neste caso teremos uma perda de dados na migração);
- O processo de transformação e limpeza de dados que ocorre durante a migração está sujeito a erros causados por má especificação de requisitos ou erros de programação;
- Os procedimentos de validação podem ser difíceis de especificar e de automatizar.

Apesar da migração de dados ser um processo arriscado e que consome alguns recursos os benefícios poderão ultrapassar os custos uma vez que o sistema legado não precisa mais de ser mantido. Atualmente a migração de dados representa um dos principais tópicos de investigação e de negócio, no entanto, há apenas algumas aproximações a este processo. Dado o elevado número de sistemas “antigos” em operação e a necessitar de migração e a variedade de problemas que levam à sua migração não é razoável pensar que se conseguirá um dia ter um processo de migração especificado que sirva para todos. A especificação a um nível abstrato é, no entanto, possível e pode servir como um guia para a implementação de um processo deste tipo [2].

Pelas razões expostas, há necessidade de uma metodologia formal na qual as organizações se possam basear para conduzir um processo de migração do seu repositório de dados. Este artigo propõe uma metodologia que poderá ser usada como guia de implementação num processo de migração dum repositório de dados legado: são descritas as fases e as atividades do processo e os riscos associados a cada uma.

### Trabalho relacionado

No contexto das TIC, o termo “*migração*” pode ter vários significados. Por exemplo e considerando apenas a preservação digital, podemos estar a falar de migração de ficheiros, migração de suporte físico, migração do formato digital, migração do repositório ou conversão/migração de dados e metadados. Cada um destes tipos de migração tem as suas particularidades muito distintas e é tratado de forma específica e diferenciada.

No contexto da preservação digital, o termo

“*repositório*” é normalmente utilizado com um dos seguintes significados:

- 1) O “repositório” é um conjunto de políticas, normas, e infraestruturas tecnológicas que constituem a plataforma de suporte à preservação digital [8], e
- 2) O “repositório” é um sistema de gestão de informação, i.e., um sistema composto por software e hardware ao qual confiamos a gestão de informação digital de acordo com certas regras [8].

No contexto deste artigo, iremos usar a segunda definição. Um repositório é um sistema composto por hardware e software que é criado para seguir certas regras e políticas e é responsável pela segurança e manutenção da informação digital que armazena.

Um repositório armazena, normalmente, vários tipos de informação, por exemplo:

- **Dados** – o recurso mais importante de todos os que são geridos no repositório; os dados representam aquilo em que os utilizadores estão interessados e que esperam que esteja bem guardado e acessível (e.g. imagens, áudio, vídeo, documentos, datasets, objetos 3D, etc.);
- **Metadados** – informação sobre os dados e que é gerida pelo repositório. É criada para suportar vários tipos de funcionalidades de valor acrescentado: pesquisa de informação, assegurar a autenticidade e proveniência dos dados, caracterização dos dados, etc.;
- **Informação específica do sistema** – informação gerada automaticamente pelo sistema e necessária ao funcionamento deste (e.g. configurações, logs, índices, informação sobre utilizadores, ficheiros de parametrização, etc.).

A “migração de um repositório” é o processo de transferência de informação digital entre dois ou mais sistemas de informação.

Na **Error! Reference source not found.**, podemos ver duas classes de migrações de repositórios. A primeira classe designada por “component migration” na qual os sistemas legados são divididos em componentes e cada componente é migrado em separado. Haverá um período de transição no qual quer o sistema legado quer o novo sistema terão de estar operacionais e a trabalhar em conjunto. Duas estratégias emergem, “phased interoperability” and “parallel operations”. Ambas precisam que os dados sejam partilhados via “database gateways”, replicados nas duas plataformas, ou separados em domínios independentes a serem migrados gradualmente para uma nova plataforma [15].

A segunda classe designada por “system migration” representa os casos em que todo o sistema legado é transferido como um todo e num único passo para a nova plataforma. Há duas subclasses nesta abordagem: “no value added” (na qual o sistema permanece praticamente na mesma, quer por emulação quer por atualização) e “value added” (na qual os sistemas adquirem novas capacidades) [15].

Esta última abordagem, “value added migration”, leva a alterações na interface com o utilizador, na base de dados e na lógica de controlo (programa). Apesar da migração ser mais complexa, esta é a abordagem que

representa um maior benefício para a organização. Pode tornar o sistema mais flexível, mais adequado às necessidades, mais fácil de manter e, consequentemente, menos dispendioso[15].

## Público Alvo

Este guia de boas práticas é dirigido a todos aqueles que trabalham ou são responsáveis por repositórios digitais de informação e que procuram orientações sobre como conduzir a migração de um sistema legado ou de um repositório que por algum motivo necessita de ser atualizado. Algumas instituições poderão mesmo usar estas recomendações quando estiverem a adquirir o serviço de migração a fornecedores externos. Este documento ajudará estas instituições a entender melhor o processo de migração de modo a poderem alocar os recursos necessários, a colocar as várias metas a atingir e a criar os processos de validação e monitorização que vão garantir a qualidade final da migração.

## Abordagem

No desenvolvimento deste guia de boas práticas, adoptou-se o seguinte procedimento dividido em quatro etapas:

### Etapla 1 – Levantamento de documentação sobre boas práticas

Pesquisou-se e juntou-se a documentação relevante produzida nos últimos anos. A pesquisa revelou que esta área de investigação está suficientemente madura e que o fenómeno da migração de repositórios está bem documentado [2, 3, 6, 9–11]. Adicionalmente, vários *whitepapers*, comunicações e relatórios técnicos provenientes da indústria das TIC foram encontrados e levados em consideração pois forneceram input valioso a este guia [1, 4, 5, 7, 11–14, 16].

### Etapla 2 – Identificação das metodologias para migração de repositórios

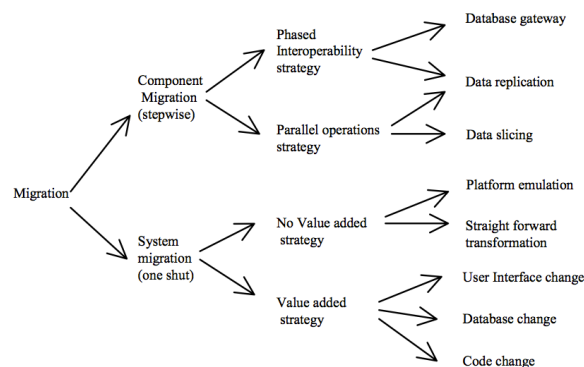
Alguns dos documentos recolhidos na fase anterior documentam metodologias para suportar a migração de sistemas de informação legados. Algumas das metodologias são simples e genéricas, outras, complexas e específicas. Esta fase consistiu na análise e avaliação desta documentação e na seleção daquela que continha informação bem alicerçada e sólida.

### Etapla 3 – Comparação de metodologias para a migração de repositórios

Algunas das metodologias identificadas são muito simples, compostas por quatro passos genéricos que podem ser aplicados a qualquer processo de desenvolvimento de software. Outras estão extremamente detalhadas. Para se poderem comparar todas, criou-se uma matriz de comparação em que todos os passos comuns estão alinhados para facilitar a comparação (Tabela 1).

### Etapla 4 – Criação de uma metodologia unificada para a migração de repositórios

A última fase na criação deste guia consistiu na classificação e generalização das atividades encontradas nas metodologias selecionadas. Este processo permitiu combinar todas as abordagens numa única metodologia unificada. Incorporaram-se também algumas alterações terminológicas de modo a adaptar o guia à área da preservação digital.



**Figura 1 - Classificação de abordagens à migração de repositórios.**

## GUIA DE BOAS PRÁTICAS PARA A MIGRAÇÃO DE REPOSITÓRIOS

Este guia corresponde, em primeiro lugar, a um exercício de formalização que tem como objetivo identificar e descrever as fases de um processo de migração de um repositório de informação digital. Não pretende ser completo ou mesmo obrigatório mas sim, fornecer informação a fornecedores, clientes e especialistas das TIC de modo a que estes possam verificar se os passos mais importantes num processo desta natureza foram seguidos.

A Figura 1 ilustra a metodologia unificada para suportar o processo de migração de um repositório. A metodologia tem dez fase cada uma delas composta por uma série de atividades.

Nas próximas secções descrevem-se estas fases e atividades.

### Análise

A primeira fase de um processo de migração tem como objetivo conhecer os sistemas, o velho e o novo, as características dos dados e as necessidades de todas as partes envolvidas para definir a melhor estratégia de migração e formalizar os requisitos. Isto inclui uma análise detalhada do sistema legado e do sistema alvo, da informação a ser migrada e das necessidades de negócio esperadas. A qualidade da análise realizada irá ditar o sucesso ou o fracasso do processo de migração.

#### 1A – Caraterização do sistema legado

O primeiro passo na migração de um repositório deve ser a análise do ambiente tecnológico do sistema legado. O processo de migração requer um levantamento completo das tecnologias envolvidas: hardware, redes, software, linguagens de programação, estruturas de dados, serviços, servidores, e requisitos temporais (e.g. disponibilidade do sistema).

Para se poder dimensionar e planificar o processo de migração, assim como orçamentar, é necessário compreender a complexidade, as relações, a qualidade, e o volume do sistema legado e dos seus dados. Isto permitirá a especificação dos requisitos apropriados tais como a a necessidade de replicação, a calendarização do projeto, tempos de resposta do sistema, fornecedores que devem ser contactados, e as configurações de

hardware necessárias para albergar os dados legados e aqueles que não-de ser gerados no futuro.

Os custos de manutenção do sistema legado também devem ser calculados. Isto permitirá clarificar as vantagens da migração e selecionar as melhores opções para a migração.

### **1B – Caracterização do sistema alvo**

Quando se planifica uma migração é importante conhecer a arquitetura e as capacidades do sistema alvo. Este conhecimento permitirá guiar o desenvolvimento das rotinas de migração, os mapeamentos estruturais, a seleção de dados, o comportamento temporal, etc.

### **1C – Análise de dados**

Um sistema legado é normalmente composto por vários tipos de informação, alguma dela estruturada e outra não estruturada. Para se migrar toda esta informação é necessário localizá-la, examiná-la e delimitá-la.

O objetivo deste passo é identificar as fontes de informação que têm de ser transportadas para o novo sistema. As fontes de informação incluem todo o tipo de informação mantida ou gerada pelo sistema legado (e.g. objetos digitais, metadados, logs, informação de utilizadores e configurações).

Há situações em que nem toda a informação deve ser preservada levando a que alguma seja deixada de fora do processo de migração. Para melhor decidir é conveniente analisar as aplicações e falar com os utilizadores das mesmas pois estes sabem qual é a informação crítica. Pode-se chegar à conclusão de que os custos da migração são proibitivos face ao volume de dados que irá ser migrado. Neste caso, um compromisso sobre que dados migrar terá de ser negociado.

Nesta fase, deve-se também fazer um levantamento das condições de acesso aos dados, dos requisitos de retenção e das medidas de segurança tais como a encriptação. Há repositórios que detêm informação muito “sensível” e com grandes restrições de segurança e acesso. Nestes casos, alguns cuidados especiais terão de ser observados de modo a proteger a informação da equipe de migração.

### **1D – Planeamento estratégico**

O objetivo do planeamento estratégico é identificar os requisitos operacionais e de negócio que poderão ter impacto na migração. As várias partes interessadas deverão ser consultadas de modo a que os seus requisitos sejam colocados no plano de migração.

Esta fase tem em consideração a informação recolhida nas fases anteriores e define a estratégia de migração a ser seguida nos passos seguintes. A estratégia de migração depende do tamanho, da complexidade e dos requisitos de negócio do sistema que se quer migrar. Por exemplo:

- Migração simples: tipicamente, corresponde à transferência de informação de um sistema para outro numa única iteração e sem grandes alterações estruturais [7, 15].
- Migração de complexidade média: pode envolver a migração de um sistema para um conjunto de novos sistemas. A migração necessita de várias iterações, poderá haver várias operações de transformação

nos dados e a integração de tudo num único modelo é normalmente complexa [7, 15].

- Migração complexa: consiste na criação de uma solução que permita a coexistência dos dois sistemas. A solução é desenhada de modo a que o sistema antigo e o novo possam cooperar e trabalhar em paralelo [7, 15].

### **1E – Especificação de requisitos**

Os passos anteriores permitem ao gestor do projeto estimar os recursos necessários para executar uma migração com sucesso. Nesta fase, o gestor do projeto definirá os critérios de sucesso do projeto: requisitos ao nível dos serviços, expectativas sobre a nova infraestrutura de armazenamento, redução de custos de manutenção, redução de custos de armazenamento, maior previsão de custos, maior flexibilidade e estabilidade técnica [4].

Os requisitos de mais alto nível também devem ser definidos neste passo: dados a migrar, requisitos de desempenho e um plano de contingência.

### **Planeamento**

Neste ponto, o gestor de projeto possui toda a informação necessária para criar o plano do projeto de migração e criar as especificações necessárias a partir das quais se poderão desenvolver as rotinas de migração e os procedimentos de teste e validação.

### **2A – Plano do projeto**

Após a análise pode-se começar a especificar o plano do projeto. O plano do projeto deverá conter a descrição da estratégia de migração que será seguida, a especificação do domínio da migração, a especificação de uma calendarização, a identificação dos recursos necessários (humanos e outros), a definição dos requisitos técnicos e de negócio, as expectativas dos clientes, a descrição dos resultados finais do projeto e um plano detalhado de execução.

Este plano, que constitui o resultado final desta fase, funcionará de guia na implementação da migração.

### **2B – Especificação das rotinas de migração**

Durante a análise já se tomaram decisões sobre quais as entidades e as fontes de dados que serão migradas. No entanto, será agora, na especificação das rotinas de migração que será criado o mapeamento entre elementos do sistema legado e os elementos do novo sistema.

Um projeto de migração representa também uma oportunidade para “se limpar e arrumar a casa”. Os detentores do repositório original são encorajados a percorrer e a analisar a informação no seu repositório identificando informação redundante ou desatualizada que pode ser eliminada, reduzindo o volume de informação sobre o qual incidirá a migração. A limpeza de dados permite aumentar a qualidade da informação armazenada mas tem os seus custos por isso há que balancear esta operação face aos benefícios que dela advirão. [13].

### **2C – Especificação do plano de testes**

Depois de todas as atividades de análise estarem concluídas, possui-se toda a informação necessária para especificar um plano de testes. Este plano deverá incluir todos os passos necessários para garantir que todos os

requisitos identificados nas fases anteriores foram satisfeitos. O plano de testes deve ser o mais completo e específico possível. Por exemplo, deverá responder a questões do tipo: O número de registos de metadados é o mesmo do sistema antigo? Todos os atributos definidos pelo utilizador foram migrados? Há problemas ao nível da codificação da informação (“*encoding*”)? Foi, algum ficheiro, corrompido durante o processo? Os invariantes do sistema antigo foram preservados no novo sistema?

O plano de testes pode implementado integralmente por *scripts*, rotinas automáticas, manualmente, ou por uma combinação destes todos. De qualquer maneira, há sempre uma verificação humana do processo pelo que, na prática, todos os planos de teste resultam de uma combinação de verificações automáticas e manuais.

O plano de testes pode ser revisto e reformulado durante as fases seguintes do processo.

### **Desenvolvimento**

Depois da análise e do planeamento, pode-se iniciar o desenvolvimento de todas as rotinas de migração e procedimentos de teste. Nesta fase, criam-se ou configuram-se as ferramentas de migração de acordo com as especificações obtidas nas fases anteriores.

#### **3A – Desenvolvimento das rotinas de migração**

Nesta fase, entra em jogo a equipe de desenvolvimento que irá desenvolver as rotinas especificadas anteriormente. Consistirá em desenvolver *scripts* de ETL (*Extract, transform and load*), programas específicos e *scripts* que implementam os mapeamentos especificados nas fases anteriores. Todos os mapeamentos, invariantes, requisitos de qualidade e validação de campos devem ser colocados nestas rotinas.

Esta fase poderá ter de ser iterada várias vezes dependendo dos erros que vão sendo detetados na validação final. Não será estranho ter de iterar de 6 a 8 vezes à volta desta fase [16].

#### **3B – Desenvolvimento dos procedimentos de teste**

Os procedimentos de teste irão validar o sucesso da migração. Os resultados desta fase podem ser listas de itens a ser verificados, *scripts* de teste ou programas dedicados que geram relatórios sobre qualquer anomalia encontrada durante o processo de migração.

#### **Inicialização do processo de migração**

Neste ponto, é necessário criar as infraestruturas onde o sistema final e a solução de migração irão operar. Para que a migração seja eficaz a infraestrutura deve ser preparada para testes globais envolvendo todo o volume de dados que se pretende migrar. Se as especificações tiverem sido bem feitas este passo será rotineiro e previsível [14].

#### **4A – Preparação do sistema final**

O sistema final é agora preparado para a transferência final de informação: preparação do hardware, sistemas operativos, configuração do armazenamento, instalação do sistema final e sua parametrização de modo a receber os dados da migração e a respeitar as regras de negócio definidas [12].

#### **4B – Ensaios e testes**

Antes da migração definitiva ser executada devem-se executar alguns ensaios para assegurar que todos os requisitos estão a ser cobertos pelas rotinas de migração.

Estes ensaios poderão incidir sobre parte dos dados ou sobre toda a informação a ser migrada. Um ensaio completo incidindo sobre toda a informação é sempre aconselhável mas poderá haver contextos que não o permitam (por exemplo a sobrecarga que isto implica sobre o sistema em produção, ou mesmo a paragem deste).

Após cada ensaio deve-se aplicar o plano de testes que irá indicar se se pode prosseguir para a migração definitiva ou se é necessário voltar atrás e alterar alguma coisa [12].

O teste final é feito colocando o novo sistema após migração à frente dos utilizadores que participaram na fase de análise e levantamento de requisitos. Estes darão o veredicto final sobre a qualidade final da migração [6].

### **Execução**

Se até aqui não ocorreram erros pode-se passar à fase de execução da migração.

É agora que todo o trabalho desenvolvido até aqui será testado. Qualquer falha que se detete irá ativar o plano de contingência previamente definido.

#### **5A – Execução das rotinas de migração**

A migração propriamente dita é realizada quando se executam as rotinas desenvolvidas em 3A.

É conveniente fazer uma última verificação antes de prosseguir: rever o plano, verificar o plano de contingência, etc.

Antes de proceder à migração é preciso eliminar do sistema novo toda a informação resultante dos ensaios e testes.

Durante a execução deve-se ter presente o plano de contingência. Uma boa estratégia é manter o sistema antigo em funcionamento ao qual poderemos recorrer se a migração correr mal.

#### **Validação**

Após a migração irá ser necessário correr uma bateria de testes no novo sistema. Estes testes irão assegurar de que tudo correu conforme planeado e não há erros gerados durante o processo. Nesta fase poderá ser gerado um relatório caracterizando os dados do novo sistema.

#### **6A – Execução das rotinas de teste**

As rotinas de teste são executadas no novo sistema. Se alguma inconsistência for detetada, pelos testes ou pelos utilizadores, o plano de contingência poderá ter de ser acionado resultando numa nova migração a ser realizada após a correção dos erros detetados.

Nalguns casos, os erros detetados poderão ser corrigidos no sistema novo, não sendo necessário proceder a uma nova migração.

#### **6B – Relatório**

A elaboração do relatório de validação corre em paralelo com os testes. A produção deste relatório é

importante pois produzirá evidências sobre a qualidade do processo, constituirá prova do sucesso do processo e poderá evitar problemas futuros de índole jurídica ou outra.

Um passo que se deve adicionar nesta fase é o arquivo das rotinas de migração juntamente com o seu relato/documentação. Apesar de serem específicas de um sistema encerram conhecimento que poderá ser reutilizado em futuras migrações.

[13].

## **6C – O fim do sistema antigo**

Assim que o novo sistema estiver pronto, com toda a informação legada migrada pode-se passar a utilizar o novo sistema. Há três estratégias para se proceder à transição [3]:

1. Desligar o sistema antigo e simplesmente começar a usar o novo;
2. Numa estratégia de interoperabilidade faseada, o sistema antigo vai-se desligando em passos incrementais, cada passo implica a substituição de alguns componentes do sistema legado;
3. Numa estratégia de paralelismo operacional, os dois sistemas são mantidos a funcionar durante um período de tempo. Durante este período o novo sistema é continuamente testado. Assim que houver confiança no seu desempenho o sistema antigo é finalmente desligado.

A primeira é arriscada pois na eventualidade de um problema não há possibilidade de voltar atrás. A segunda pode revelar-se extremamente complexa além de que os sistemas teriam de estar muito bem estruturados o que não é normalmente o caso. A terceira é que pode dar mais garantias mas como é normal em situações semelhantes o ideal poderá passar por um misto das três estratégias coisa que tem de ser feita caso a caso [3].

## **Finalização**

Depois do novo repositório entrar em produção há uma série de atividades que devem ser consideradas: dar formação aos utilizadores no novo sistema, juntar e arquivar toda a documentação gerada bem como todas as rotinas desenvolvidas e providenciar uma linha de suporte e manutenção para o novo sistema.

## **7A – Formação**

A adopção de um sistema só ficará completa depois dos seus utilizadores receberem formação na sua utilização. O feedback recebido dos utilizadores durante estas ações de formação é muito importante e poderá ajudar a parametrizar melhor o novo sistema bem dar informação preciosa sobre a condução de futuras migrações

A formação deve ser dada no fim do projeto porque o seu tempo de vida útil é curto se os utilizadores não colocarem em prática de imediato os novos conhecimentos. Poderão, no entanto, ser consideradas ações de formação mais cedo no projeto com o objetivo de preparar alguns utilizadores para ajudarem na configuração do novo sistema [14].

## **7B – Documentação**

Depois de terminada a migração, a equipe de projeto deve reunir todo o material documental produzido, estatísticas, relatórios, planos, especificações de design e preparar um relatório final onde deverão ser realçadas as atividades que correram bem e as que correram mal. Este relatório deve ser partilhado com toda a equipe com o objetivo de aumentar o conhecimento de cada indivíduo e dessa forma conseguir melhorias em futuros processos deste tipo [7].

## **7C - Suporte**

Esta fase, consiste em ter uma equipe de técnicos disponível para ajudar os utilizadores em questões que possam ter ou dificuldades operacionais que estejam a sentir.

## **CONCLUSÕES E TRABALHO FUTURO**

A migração de um repositório é um processo inevitável para todas as instituições que alberguem ou que estejam a gerir um repositório digital. É apenas uma questão de tempo até o repositório se tornar incompatível com as tecnologias atuais, inadequado para servir as necessidades de negócio da instituição ou simplesmente antiquado e deixando de satisfazer as expectativas da sua comunidade de utilizadores.

Apesar de ser um processo complexo e arriscado, a migração do repositório, com a preparação adequada, é definitivamente uma empreitada que vale o investimento.

Uma metodologia clara e compreensível, uma equipe bem preparada e um plano de projeto bem definido são as chaves para o sucesso. No entanto, haverá sempre lugar a falhas.

Para prevenir estes problemas, as instituições devem seguir uma metodologia consistente e confiável que lhes permita analisar, planejar, especificar, desenvolver, migrar e validar a migração. Riscos potenciais poderão ser evitados se a metodologia descrita neste documento for seguida.

Como trabalho futuro, estas recomendações serão incrementadas com recomendações sobre a implementação e exemplos reais de boas práticas. Este trabalho será então publicado como um *whitepaper* técnico e disponibilizado em regime de acesso livre.

## **AGRADECIMENTOS**

Parte deste trabalho foi suportada pela União Europeia ao abrigo do 7º programa quadro através do projeto SCAPE - Contrato 270137.

Part of this work was supported by the European Union in the 7th Framework Program, IST, through the SCAPE project, Contract 270137.

## **REFERÊNCIAS**

- [1] Bearing Point Inc. 2008. Data Migration through an Information Development Approach - A management overview.
- [2] Bisbal, J. et al. 1999. Legacy information systems: issues and directions. IEEE Software. 16, 5 (1999), 103-111.
- [3] Brodie, M.L. and Stonebraker, M. 1995. Migrating legacy systems: gateways, interfaces & the incremental approach. Morgan Kaufmann Publishers Inc.

- [4] Burry, C. and Mancusi, D. 2004. How to plan for data migration. Computer World.
- [5] Harris, L. 2010. IPM 11g Migration Best Practices. Nexus'10 (2010), 27.
- [6] Hudicka, J.R. 1998. An Overview of Data Migration Methodology. Select Magazine. (1998), 5.
- [7] IBM Global Technology Services 2007. Best practices for data migration - Methodologies for planning, designing, migrating and validating data migration.
- [8] Jantz, R. 2005. Digital Preservation: Architecture and Technology for Trusted Digital Repositories. DLib Magazine. (2005), 1-17.
- [9] Lawless, D. et al. 1997. The Butterfly Methodology : A Gateway-free Approach for Migrating Legacy Information Systems. Proceedings Third IEEE International Conference on Engineering of Complex Computer Systems Cat No97TB100168. (1997), 200-205.
- [10] De Lucia, A. et al. 2008. Developing legacy system migration methods and tools for technology transfer. Software Practice and Experience. 38, 13 (2008), 1333-1364.
- [11] Mohanty, S. 2004. Data Migration Strategies, Part 1. Information Management.
- [12] Network Appliance Inc. 2006. Data migration best practices.
- [13] Open Text Corporation 2009. Top 10 best practices in content migration.
- [14] Pick, B.G. 2001. Data Migration Concepts & Challenges.
- [15] Rahgozar, M. and Oroumchian, F. 2003. An effective strategy for legacy systems evolution. Journal of Software Maintenance and Evolution Research and Practice. 15, 5 (2003), 325-344.
- [16] Utopia Inc. 2009. Data migration management - A methodology: Sustaining data integrity after the go live and beyond.

## APÊNDICES

**Tabela 1 - Matriz de comparação de metodologias para migração de sistemas legados.**

		System migration methodologies													
Activity	Sub activity	[5]	[6]	[14]	[1]	[4]	[11]	[12]	[7]	[16]	[13]	[10]	[9]	[3]	
Analysis & consultation	Characterization of legacy environment		✓			✓			✓	✓	✓	✓		✓	
	Characterization of target environment										✓	✓	✓		
	Data analysis		✓			✓		✓		✓	✓	✓	✓	✓	
	Strategic planning		✓		✓			✓		✓	✓				
	Definition of requirements	✓		✓	✓	✓			✓				✓		
Planning & design	Project planning	✓		✓	✓		✓	✓	✓	✓					
	Design of migration routines	✓	✓		✓	✓	✓	✓	✓		✓	✓	✓	✓	
	Design test plan								✓		✓				
Development	Development of migration routines	✓		✓	✓				✓	✓		✓	✓		
	Development of testing routines														
Setup and testing	Target environment provisioning				✓			✓	✓			✓	✓	✓	
	Rehearsal & testing		✓				✓	✓	✓	✓		✓	✓		
Execute	Execution of migration routines	✓		✓	✓		✓	✓	✓	✓					
Validate	Execution of testing routines	✓	✓		✓	✓	✓	✓	✓		✓	✓	✓	✓	
	Reporting								✓		✓				
	Cut-over	✓		✓	✓				✓	✓		✓	✓		
Wrap up	Training														
	Documenting				✓			✓	✓			✓	✓	✓	
	Supporting		✓				✓	✓	✓	✓		✓	✓		